

SEQUENCE LISTING

<110> Schweifer, Norbert
Scherl-Mostageer, Marwa
Sommergruber, Wolfgang
Abseher, Roger

<120> Tumorassoziiertes Antigen (B345)

<130> 0652.2280001/EKS/AES

<140>

<141>

<150> DE 100 33 080.0

<151> 2000-07-07

<150> DE 101 19 294.0

<151> 2001-04-19

<150> US 60/243,158

<151> 2000-10-25

<150> US 60/297,747

<151> 2001-06-14

<160> 40

<170> PatentIn Ver. 2.1

<210> 1

<211> 5897

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> (1)..(214)

<220>

<221> CDS

<222> (215)..(2464)

<220>

<221> 3'UTR

<222> (2465)..(5897)

<400> 1

cttgagatat tagaattcgc gactcctgaa ctgcgggggtc tctatcgcac tgctaggggt 60

tctgctgctg ggtgcggcgc gcctgccgcg cggggcagaa gcttttgaga ttgctctgcc 120

acgagaaagc aacattacag ttctcataaa gctggggacc ccgactctgc tggcaaaacc 180

ctgttacatc gtcattttcta aaagacatat aacc atg ttg tcc atc aag tct gga 235

Met Leu Ser Ile Lys Ser Gly

1

5

gaa aga ata gtc ttt acc ttt agc tgc cag agt cct gag aat cac ttt 283

Glu Arg Ile Val Phe Thr Phe Ser Cys Gln Ser Pro Glu Asn His Phe

10

15

20

gtc ata gag atc cag aaa aat att gac tgt atg tca ggc cca tgt cct 331

Val Ile Glu Ile Gln Lys Asn Ile Asp Cys Met Ser Gly Pro Cys Pro

25

30

35

ttt ggg gag gtt cag ctt cag ccc tcg aca tcg ttg ttg cct acc ctc 379

Phe Gly Glu Val Gln Leu Gln Pro Ser Thr Ser Leu Leu Pro Thr Leu

40

45

50

55

aac aga act ttc atc tgg gat gtc aaa gct cat aag agc atc ggt tta 427

Asn Arg Thr Phe Ile Trp Asp Val Lys Ala His Lys Ser Ile Gly Leu

60

65

70

gag ctg cag ttt tcc atc cct cgc ctg agg cag atc ggt ccg ggt gag 475

Glu Leu Gln Phe Ser Ile Pro Arg Leu Arg Gln Ile Gly Pro Gly Glu

75

80

85

agc tgc cca gac gga gtc act cac tcc atc agc ggc cga atc gat gcc 523

Ser Cys Pro Asp Gly Val Thr His Ser Ile Ser Gly Arg Ile Asp Ala

90

95

100

acc gtg gtc agg atc gga acc ttc tgc agc aat ggc act gtg tcc cgg 571

Thr Val Val Arg Ile Gly Thr Phe Cys Ser Asn Gly Thr Val Ser Arg

105

110

115

atc aag atg caa gaa gga gtg aaa atg gcc tta cac ctc cca tgg ttc 619

Ile Lys Met Gln Glu Gly Val Lys Met Ala Leu His Leu Pro Trp Phe

120

125

130

135

cac ccc aga aat gtc tcc ggc ttc agc att gca aac cgc tca tct ata 667

His Pro Arg Asn Val Ser Gly Phe Ser Ile Ala Asn Arg Ser Ser Ile

140

145

150

aaa cgt ctg tgc atc atc gag tct gtg ttt gag ggt gaa ggc tca gca 715

Lys Arg Leu Cys Ile Ile Glu Ser Val Phe Glu Gly Glu Gly Ser Ala

155

160

165

acc ctg atg tct gcc aac tac cca gaa ggc ttc cct gag gat gag ctc 763

Thr Leu Met Ser Ala Asn Tyr Pro Glu Gly Phe Pro Glu Asp Glu Leu

170

175

180

atg acg tgg cag ttt gtc gtt cct gca cac ctg cgg gcc agc gtc tcc 811

Met Thr Trp Gln Phe Val Val Pro Ala His Leu Arg Ala Ser Val Ser

185	190	195	
ttc ctc aac ttc aac ctc tcc aac tgt gag agg aag gag gag cgg gtt			859
Phe Leu Asn Phe Asn Leu Ser Asn Cys Glu Arg Lys Glu Glu Arg Val			
200	205	210	215
gaa tac tac atc ccg ggc tcc acc acc aac ccc gag gtg ttc aag ctg			907
Glu Tyr Tyr Ile Pro Gly Ser Thr Thr Asn Pro Glu Val Phe Lys Leu			
	220	225	230
gag gac aag cag cct ggg aac atg gcg ggg aac ttc aac ctc tct ctg			955
Glu Asp Lys Gln Pro Gly Asn Met Ala Gly Asn Phe Asn Leu Ser Leu			
	235	240	245
caa ggc tgt gac caa gat gcc caa agt cca ggg atc ctc cgg ctg cag			1003
Gln Gly Cys Asp Gln Asp Ala Gln Ser Pro Gly Ile Leu Arg Leu Gln			
	250	255	260
ttc caa gtt ttg gtc caa cat cca caa aat gaa agc aat aaa atc tac			1051
Phe Gln Val Leu Val Gln His Pro Gln Asn Glu Ser Asn Lys Ile Tyr			
	265	270	275
gtg gtt gac ttg agt aat gag cga gcc atg tca ctc acc atc gag cca			1099
Val Val Asp Leu Ser Asn Glu Arg Ala Met Ser Leu Thr Ile Glu Pro			
280	285	290	295
cgg ccc gtc aaa cag agc cgc aag ttt gtc cct ggc tgt ttc gtg tgt			1147
Arg Pro Val Lys Gln Ser Arg Lys Phe Val Pro Gly Cys Phe Val Cys			
	300	305	310
cta gaa tct cgg acc tgc agt agc aac ctc acc ctg aca tct ggc tcc			1195
Leu Glu Ser Arg Thr Cys Ser Ser Asn Leu Thr Leu Thr Ser Gly Ser			
	315	320	325
aaa cac aaa atc tcc ttc ctt tgt gat gat ctg aca cgt ctg tgg atg			1243
Lys His Lys Ile Ser Phe Leu Cys Asp Asp Leu Thr Arg Leu Trp Met			
	330	335	340
aat gtg gaa aaa acc ata agc tgc aca gac cac cgg tac tgc caa agg			1291
Asn Val Glu Lys Thr Ile Ser Cys Thr Asp His Arg Tyr Cys Gln Arg			
	345	350	355
aaa tcc tac tca ctc cag gtg ccc agt gac atc ctc cac ctg cct gtg			1339
Lys Ser Tyr Ser Leu Gln Val Pro Ser Asp Ile Leu His Leu Pro Val			
360	365	370	375
gag ctg cat gac ttc tcc tgg aag ctg ctg gtg ccc aag gac agg ctc			1387
Glu Leu His Asp Phe Ser Trp Lys Leu Leu Val Pro Lys Asp Arg Leu			
	380	385	390
agc ctg gtg ctg gtg cca gcc cag aag ctg cag cag cat aca cac gag			1435

Ser	Leu	Val	Leu	Val	Pro	Ala	Gln	Lys	Leu	Gln	Gln	His	Thr	His	Glu	
			395					400					405			
aag	ccc	tgc	aac	acc	agc	ttc	agc	tac	ctc	gtg	gcc	agt	gcc	ata	ccc	1483
Lys	Pro	Cys	Asn	Thr	Ser	Phe	Ser	Tyr	Leu	Val	Ala	Ser	Ala	Ile	Pro	
		410					415					420				
agc	cag	gac	ctg	tac	ttc	ggc	tcc	ttc	tgc	ccg	gga	ggc	tct	atc	aag	1531
Ser	Gln	Asp	Leu	Tyr	Phe	Gly	Ser	Phe	Cys	Pro	Gly	Gly	Ser	Ile	Lys	
		425				430					435					
cag	atc	cag	gtg	aag	cag	aac	atc	tcg	gtg	acc	ctt	cgc	acc	ttt	gcc	1579
Gln	Ile	Gln	Val	Lys	Gln	Asn	Ile	Ser	Val	Thr	Leu	Arg	Thr	Phe	Ala	
440					445					450					455	
ccc	agc	ttc	caa	caa	gag	gcc	tcc	agg	cag	ggc	ctg	acg	gtg	tcc	ttt	1627
Pro	Ser	Phe	Gln	Gln	Glu	Ala	Ser	Arg	Gln	Gly	Leu	Thr	Val	Ser	Phe	
			460					465						470		
ata	cct	tat	ttc	aaa	gag	gaa	ggc	gtt	ttc	acg	gtg	acc	cct	gac	aca	1675
Ile	Pro	Tyr	Phe	Lys	Glu	Glu	Gly	Val	Phe	Thr	Val	Thr	Pro	Asp	Thr	
			475				480						485			
aaa	agc	aag	gtc	tac	ctg	agg	acc	ccc	aac	tgg	gac	cgg	ggc	ctg	cca	1723
Lys	Ser	Lys	Val	Tyr	Leu	Arg	Thr	Pro	Asn	Trp	Asp	Arg	Gly	Leu	Pro	
		490				495					500					
tcc	ctc	acc	tct	gtg	tcc	tgg	aac	atc	agc	gtg	ccc	aga	gac	cag	gtg	1771
Ser	Leu	Thr	Ser	Val	Ser	Trp	Asn	Ile	Ser	Val	Pro	Arg	Asp	Gln	Val	
	505					510					515					
gcc	tgc	ctg	act	ttc	ttt	aag	gag	cgg	agc	ggc	gtg	gtc	tgc	cag	aca	1819
Ala	Cys	Leu	Thr	Phe	Phe	Lys	Glu	Arg	Ser	Gly	Val	Val	Cys	Gln	Thr	
520					525					530					535	
ggg	cgc	gca	ttc	atg	atc	atc	cag	gag	cag	cgg	acc	cgg	gct	gag	gag	1867
Gly	Arg	Ala	Phe	Met	Ile	Ile	Gln	Glu	Gln	Arg	Thr	Arg	Ala	Glu	Glu	
			540				545						550			
atc	ttc	agc	ctg	gac	gag	gat	gtg	ctc	ccc	aag	cca	agc	ttc	cac	cat	1915
Ile	Phe	Ser	Leu	Asp	Glu	Asp	Val	Leu	Pro	Lys	Pro	Ser	Phe	His	His	
			555				560						565			
cac	agc	ttc	tgg	gtc	aac	atc	tct	aac	tgc	agc	ccc	acg	agc	ggc	aag	1963
His	Ser	Phe	Trp	Val	Asn	Ile	Ser	Asn	Cys	Ser	Pro	Thr	Ser	Gly	Lys	
		570				575					580					
cag	cta	gac	ctg	ctc	ttc	tcg	gtg	aca	ctt	acc	cca	agg	act	gtg	gac	2011
Gln	Leu	Asp	Leu	Leu	Phe	Ser	Val	Thr	Leu	Thr	Pro	Arg	Thr	Val	Asp	
		585				590					595					
ttg	act	gtc	atc	ctc	atc	gca	gcg	gtg	gga	ggc	gga	gtc	tta	ctg	ctg	2059
Leu	Thr	Val	Ile	Leu	Ile	Ala	Ala	Val	Gly	Gly	Gly	Val	Leu	Leu	Leu	

600	605	610	615	
tct gcc ctc ggg ctc atc att tgc tgt gtg aaa aag aag aaa aag aag	2107			
Ser Ala Leu Gly Leu Ile Ile Cys Cys Val Lys Lys Lys Lys Lys Lys				
620	625	630		
aca aac aag ggc ccc gct gtg ggt atc tac aat ggc aac atc aat act	2155			
Thr Asn Lys Gly Pro Ala Val Gly Ile Tyr Asn Gly Asn Ile Asn Thr				
635	640	645		
gag atg cca ggc agc caa aaa agt ttc aga aag ggc gaa agg aca atg	2203			
Glu Met Pro Gly Ser Gln Lys Ser Phe Arg Lys Gly Glu Arg Thr Met				
650	655	660		
act ccc atg tgt atg cag tca tgc agg aca cca tgg tat atg ggc atc	2251			
Thr Pro Met Cys Met Gln Ser Ser Arg Thr Pro Trp Tyr Met Gly Ile				
665	670	675		
tgc tac agg att cca gcg gct cct tcc tgc agc cag agg tgg aca cct	2299			
Cys Tyr Arg Ile Pro Ala Ala Pro Ser Cys Ser Gln Arg Trp Thr Pro				
680	685	690	695	
acc ggc cgt tcc agg gca cca tgg ggg tct gtc ctc cct ccc cac cca	2347			
Thr Gly Arg Ser Arg Ala Pro Trp Gly Ser Val Leu Pro Pro His Pro				
700	705	710		
cca tat gct cca ggg ccc caa ctg caa agt tgg cca ctg agg agc cac	2395			
Pro Tyr Ala Pro Gly Pro Gln Leu Gln Ser Trp Pro Leu Arg Ser His				
715	720	725		
ctc ctc gct ccc ctc ctg agt ctg aga gtg aac cgt aca cct tct ccc	2443			
Leu Leu Ala Pro Leu Leu Ser Leu Arg Val Asn Arg Thr Pro Ser Pro				
730	735	740		
atc cca aca atg ggg atg taa gcagcaagga cacagacatt cccttactga	2494			
Ile Pro Thr Met Gly Met				
745	750			
acactcagga gcccatggag ccagcagaat aacttgatcc attccagacg ctttgctgag	2554			
tttcataaag cagggcactg agacacccgt ccgtgttcct aaccagaaat cctaaagaag	2614			
aggaattata cagaaggaac agcaggaggt tttcctggac accgccaact tcacattgct	2674			
cagtggactc attctaaggg caagacattg aaaatgatga attccaatct ggatacagtc	2734			
atgacagctc atgtgctcct caacttaggc tgtgcgggta gccagcctgt aatgagagga	2794			
gagaggcctg agtcacctag cataggggtg cagcaagccc tggattcaga gtgttaaaca	2854			
gaggcttgcc ctcttcagga caacagttcc aattccaagg agcctacctg aggtccctac	2914			

tctcactggg gtccccagga tgaaaacgac aatgtgcctt tttattatta tttatttggg 2974
ggtcctgtgt tatttaagag atcaaagtga taaccaccta gctcttttca cctgacttag 3034
taataactca tactaactgg tttggatgcc tgggttggtga cttctactga ccgctagata 3094
aacgtgtgcc tgtccccag gtggtgggaa taatttacia tctgtccaac cagaaaagaa 3154
tgtgtgtgtt tgagcagcat tgacacatat ctgctttgat aagagacttc ctgattctct 3214
aggtcgggttc gtggttatcc cattgtggaa attcatcttg aatcccattg tcctatagtc 3274
ctagcaataa gagaaatttc ctcaagtttc catgtgcggt tctcctagct gcagcaatac 3334
tttgacattt aaagagaaat ttagagaata ttctcatcct ctaaaaatgt ttaaataatat 3394
accaaacagt ggccccctgc attagttttc tgttgccact gcaaccctatt acttggttagc 3454
ttaaaaacia cacattagct tatagtcctg gggatcagaa ttccaaaatg gatgtccctg 3514
aatgaaaatc aagggtgtcag cagagctgtg ctccctctga aggtctctagg gagaagccgg 3574
ttccttgcca tttcaagctt ctagaggctg gctgcattcc caggctccag tggctgggtca 3634
agcttttctc acatggcatc actgtgacac tggccctccc acttccctct ttgacttaca 3694
aagcccacca ggaagatcca ggataatctc tccatctaaa gatccttcat catcctggaa 3754
gagccttttg ccatgcaaga caacatagcc acaggtgggg attaggacca ggacatcttt 3814
ggggtgctgt tattctgcct accacacctt cctgccacbg actcccacag gagaggctac 3874
aaaatgatct ggcgcacagg gatgttttgt ttagcttgcg gactctaaca cttaaaaaaa 3934
ccccagatca gaagatctgg ccatgctggg gctcacattc tcacctagca acaactgggt 3994
ggagctgggc accagctctg cctttagaag ggggtgtccac ttcaccaggt caccacagcc 4054
cacactacgc cctatcactt cccacaatga ggctaagtgt ttgtttctac tgatcaatgc 4114
ccctgcaggt tgcatttatt gtaatgaaaa agaaagactg ggattaatct ctaatcaggt 4174
gagtagacca tgagaccaat gtgtgtcac attacccttt ttcttttttt tctttttctt 4234
tttctttttt tttttaatgt gagacaggat ctcatctgtg tgcctaggct ggagtgcagt 4294
ggcgcaatct cggctcactg caacctctgc ctccctgggt caagcaattc tcccacctca 4354
gcctccaaa tagctgggat cactggcaca aaccaccatg cccagctaatt tttgtatttt 4414
ttgtagagac agggtttcac catgttgccc aggetgggtct caacctcctg ggctcaagca 4474
atcctcctgc ctgggcctcc caaagtgctg ggattacaga tgtgagccac cgcacccagc 4534

cccacaccct catttatacc aattacctgc ccagtaactg tggacttttg cttoctcacc 4594
 cctgctctga tctggaagga gagggattat gttatagctt gtcagcacag tcccaagttc 4654
 aatattttctg cggcaaaaac ttccttcaaa aaataaatgt acttcattgt attcaatgaa 4714
 ttcaccttgg aaatgcaccg cctcaacttg ttcacatggc ataaatgaaa ggaattttat 4774
 agtctcctaa atggcgtgta ctgcaagacc tcttgaacac tttccagagg ataggatatt 4834
 taagtcatgc ccttggcggt gcctatggca cctttccctt ctgaaagtct ggttcctgcc 4894
 cagtgaccct tggccttgtg agccgagatg ctgaccctgc ataaagggcc aaaggagggc 4954
 tgcggcttcc ttcctcact gaagagccct tatttgaatt cactgtgtgg agccctagcc 5014
 ctccattctc gacattcccc aacctccag ccccttccaa gcaggactag gtgccctgca 5074
 ttccacccaa ggtgggattg gccttcctta ggctggctac ttgtcaccat caccgacatc 5134
 actgttgctt gcaaggacac cagtgggcca ttttcttca actgagggct caaaactcct 5194
 ggacaagttg ctggctcctg agaccagtat ttcctggagm tgtgcctcag tgaaggggcc 5254
 cagcctgagg aacctgggt cttttcttta aagcccaggc cccacttaca taaaacattt 5314
 cagggctcact ggaaacagtg aagtgccatt tgtngaagcc tactgnatgc cagccactg 5374
 ctcatccacg tggatatgca tgccctacgag gaaggccagc gcatgcagga ntgggtctcta 5434
 atgntgtggt cattgcacag aagggaaagg tctcaaggaa gagtcaactg ggacaagcac 5494
 aagcccaccg gacatggcct tggtaaaggt tagcagactg gtgtgtgtgg atctgcagtg 5554
 cttcactgga aataatttat tcattgcaga tacttttttag gtggcatttt attcatttcc 5614
 tgtgctttta ataaacaaat gtaccaaaaa acaagtatca agctgtttta gtgcttcggc 5674
 tacttgtccc ctggttcagt agaggccccg gtttcccagt tgttgactgt gacaggctca 5734
 gcatgggctc agcagatgct gtcttaattt gtggatgata cagaaagcca ggctttggga 5794
 tacaagttct ttcctcttca tttgatgccg tgcactgtgt gaagcagatg tttttgtccg 5854
 gaaataaaaa taatagtctt ggagtctcgc caaaaaaaaa aag 5897

<210> 2
 <211> 749
 <212> PRT
 <213> Homo sapiens

 <400> 2

Met	Leu	Ser	Ile	Lys	Ser	Gly	Glu	Arg	Ile	Val	Phe	Thr	Phe	Ser	Cys	1	5	10	15
Gln	Ser	Pro	Glu	Asn	His	Phe	Val	Ile	Glu	Ile	Gln	Lys	Asn	Ile	Asp	20	25	30	
Cys	Met	Ser	Gly	Pro	Cys	Pro	Phe	Gly	Glu	Val	Gln	Leu	Gln	Pro	Ser	35	40	45	
Thr	Ser	Leu	Leu	Pro	Thr	Leu	Asn	Arg	Thr	Phe	Ile	Trp	Asp	Val	Lys	50	55	60	
Ala	His	Lys	Ser	Ile	Gly	Leu	Glu	Leu	Gln	Phe	Ser	Ile	Pro	Arg	Leu	65	70	75	80
Arg	Gln	Ile	Gly	Pro	Gly	Glu	Ser	Cys	Pro	Asp	Gly	Val	Thr	His	Ser	85	90	95	
Ile	Ser	Gly	Arg	Ile	Asp	Ala	Thr	Val	Val	Arg	Ile	Gly	Thr	Phe	Cys	100	105	110	
Ser	Asn	Gly	Thr	Val	Ser	Arg	Ile	Lys	Met	Gln	Glu	Gly	Val	Lys	Met	115	120	125	
Ala	Leu	His	Leu	Pro	Trp	Phe	His	Pro	Arg	Asn	Val	Ser	Gly	Phe	Ser	130	135	140	
Ile	Ala	Asn	Arg	Ser	Ser	Ile	Lys	Arg	Leu	Cys	Ile	Ile	Glu	Ser	Val	145	150	155	160
Phe	Glu	Gly	Glu	Gly	Ser	Ala	Thr	Leu	Met	Ser	Ala	Asn	Tyr	Pro	Glu	165	170	175	
Gly	Phe	Pro	Glu	Asp	Glu	Leu	Met	Thr	Trp	Gln	Phe	Val	Val	Pro	Ala	180	185	190	
His	Leu	Arg	Ala	Ser	Val	Ser	Phe	Leu	Asn	Phe	Asn	Leu	Ser	Asn	Cys	195	200	205	
Glu	Arg	Lys	Glu	Glu	Arg	Val	Glu	Tyr	Tyr	Ile	Pro	Gly	Ser	Thr	Thr	210	215	220	
Asn	Pro	Glu	Val	Phe	Lys	Leu	Glu	Asp	Lys	Gln	Pro	Gly	Asn	Met	Ala	225	230	235	240
Gly	Asn	Phe	Asn	Leu	Ser	Leu	Gln	Gly	Cys	Asp	Gln	Asp	Ala	Gln	Ser	245	250	255	
Pro	Gly	Ile	Leu	Arg	Leu	Gln	Phe	Gln	Val	Leu	Val	Gln	His	Pro	Gln	260	265	270	
Asn	Glu	Ser	Asn	Lys	Ile	Tyr	Val	Val	Asp	Leu	Ser	Asn	Glu	Arg	Ala	275	280	285	

Met	Ser	Leu	Thr	Ile	Glu	Pro	Arg	Pro	Val	Lys	Gln	Ser	Arg	Lys	Phe	
290						295		300								
Val	Pro	Gly	Cys	Phe	Val	Cys	Leu	Glu	Ser	Arg	Thr	Cys	Ser	Ser	Asn	
305		310				315								320		
Leu	Thr	Leu	Thr	Ser	Gly	Ser	Lys	His	Lys	Ile	Ser	Phe	Leu	Cys	Asp	
				325		330								335		
Asp	Leu	Thr	Arg	Leu	Trp	Met	Asn	Val	Glu	Lys	Thr	Ile	Ser	Cys	Thr	
			340		345								350			
Asp	His	Arg	Tyr	Cys	Gln	Arg	Lys	Ser	Tyr	Ser	Leu	Gln	Val	Pro	Ser	
		355		360								365				
Asp	Ile	Leu	His	Leu	Pro	Val	Glu	Leu	His	Asp	Phe	Ser	Trp	Lys	Leu	
370						375		380								
Leu	Val	Pro	Lys	Asp	Arg	Leu	Ser	Leu	Val	Leu	Val	Pro	Ala	Gln	Lys	
385		390				395								400		
Leu	Gln	Gln	His	Thr	His	Glu	Lys	Pro	Cys	Asn	Thr	Ser	Phe	Ser	Tyr	
				405		410								415		
Leu	Val	Ala	Ser	Ala	Ile	Pro	Ser	Gln	Asp	Leu	Tyr	Phe	Gly	Ser	Phe	
			420		425								430			
Cys	Pro	Gly	Gly	Ser	Ile	Lys	Gln	Ile	Gln	Val	Lys	Gln	Asn	Ile	Ser	
		435		440								445				
Val	Thr	Leu	Arg	Thr	Phe	Ala	Pro	Ser	Phe	Gln	Gln	Glu	Ala	Ser	Arg	
450						455		460								
Gln	Gly	Leu	Thr	Val	Ser	Phe	Ile	Pro	Tyr	Phe	Lys	Glu	Glu	Gly	Val	
465		470				475								480		
Phe	Thr	Val	Thr	Pro	Asp	Thr	Lys	Ser	Lys	Val	Tyr	Leu	Arg	Thr	Pro	
				485		490								495		
Asn	Trp	Asp	Arg	Gly	Leu	Pro	Ser	Leu	Thr	Ser	Val	Ser	Trp	Asn	Ile	
			500		505								510			
Ser	Val	Pro	Arg	Asp	Gln	Val	Ala	Cys	Leu	Thr	Phe	Phe	Lys	Glu	Arg	
		515		520								525				
Ser	Gly	Val	Val	Cys	Gln	Thr	Gly	Arg	Ala	Phe	Met	Ile	Ile	Gln	Glu	
530						535		540								
Gln	Arg	Thr	Arg	Ala	Glu	Glu	Ile	Phe	Ser	Leu	Asp	Glu	Asp	Val	Leu	
545		550				555								560		
Pro	Lys	Pro	Ser	Phe	His	His	His	Ser	Phe	Trp	Val	Asn	Ile	Ser	Asn	
				565		570								575		

```

Cys Ser Pro Thr Ser Gly Lys Gln Leu Asp Leu Leu Phe Ser Val Thr
      580                      585                      590

Leu Thr Pro Arg Thr Val Asp Leu Thr Val Ile Leu Ile Ala Ala Val
      595                      600                      605

Gly Gly Gly Val Leu Leu Leu Ser Ala Leu Gly Leu Ile Ile Cys Cys
      610                      615                      620

Val Lys Lys Lys Lys Lys Lys Thr Asn Lys Gly Pro Ala Val Gly Ile
      625                      630                      635                      640

Tyr Asn Gly Asn Ile Asn Thr Glu Met Pro Gly Ser Gln Lys Ser Phe
      645                      650                      655

Arg Lys Gly Glu Arg Thr Met Thr Pro Met Cys Met Gln Ser Ser Arg
      660                      665                      670

Thr Pro Trp Tyr Met Gly Ile Cys Tyr Arg Ile Pro Ala Ala Pro Ser
      675                      680                      685

Cys Ser Gln Arg Trp Thr Pro Thr Gly Arg Ser Arg Ala Pro Trp Gly
      690                      695                      700

Ser Val Leu Pro Pro His Pro Pro Tyr Ala Pro Gly Pro Gln Leu Gln
      705                      710                      715                      720

Ser Trp Pro Leu Arg Ser His Leu Leu Ala Pro Leu Leu Ser Leu Arg
      725                      730                      735

Val Asn Arg Thr Pro Ser Pro Ile Pro Thr Met Gly Met
      740                      745

```

```

<210> 3
<211> 6163
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> 5'UTR
<222> (1)..(282)

```

```

<220>
<221> GC_signal
<222> (147)..(157)

```

```

<220>
<221> misc_feature
<222> (201)..(209)
<223> cap signal; Transkriptionsstart

```

```

<220>

```

<221> 3'UTR
<222> (2794)..(6163)

<220>
<221> 3'UTR
<222> (2794)..(6163)

<220>
<221> CDS
<222> (283)..(2793)

<400> 3
ccaacgccgc aatgggggagt agtagggacc cagcaacccg gtgccgggag ccctgcaccc 60
tgggaggggag aggcggtcgc tgaggcagga agaggaggag gagagagagg agggacgcac 120
cgggtcagct cgcgacccctg ctgcgcaggg cgggggctcgg gccgggtccgc ccgcgcgcag 180
gtgagtgagc cagggcggag cgcagctgcg cggggcttgg gcgcctgggg ccgccgctcc 240
ccaccgtcgt tttccccacc gagggcggagg cgtcccggag tc atg gcc ggc ctg 294
Met Ala Gly Leu
1
aac tgc ggg gtc tct atc gca ctg cta ggg gtt ctg ctg ctg ggt gcg 342
Asn Cys Gly Val Ser Ile Ala Leu Leu Gly Val Leu Leu Leu Gly Ala
5 10 15 20
gcg cgc ctg ccg cgc ggg gca gaa gct ttt gag att gct ctg cca cga 390
Ala Arg Leu Pro Arg Gly Ala Glu Ala Phe Glu Ile Ala Leu Pro Arg
25 30 35
gaa agc aac att aca gtt ctc ata aag ctg ggg acc ccg act ctg ctg 438
Glu Ser Asn Ile Thr Val Leu Ile Lys Leu Gly Thr Pro Thr Leu Leu
40 45 50
gca aaa ccc tgt tac atc gtc att tct aaa aga cat ata acc atg ttg 486
Ala Lys Pro Cys Tyr Ile Val Ile Ser Lys Arg His Ile Thr Met Leu
55 60 65
tcc atc aag tct gga gaa aga ata gtc ttt acc ttt agc tgc cag agt 534
Ser Ile Lys Ser Gly Glu Arg Ile Val Phe Thr Phe Ser Cys Gln Ser
70 75 80
cct gag aat cac ttt gtc ata gag atc cag aaa aat att gac tgt atg 582
Pro Glu Asn His Phe Val Ile Glu Ile Gln Lys Asn Ile Asp Cys Met
85 90 95 100
tca ggc cca tgt cct ttt ggg gag gtt cag ctt cag ccc tcg aca tcg 630
Ser Gly Pro Cys Pro Phe Gly Glu Val Gln Leu Gln Pro Ser Thr Ser
105 110 115
ttg ttg cct acc ctc aac aga act ttc atc tgg gat gtc aaa gct cat 678
Leu Leu Pro Thr Leu Asn Arg Thr Phe Ile Trp Asp Val Lys Ala His

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

120	125	130	
aag agc atc ggt tta gag ctg cag ttt tcc atc cct cgc ctg agg cag			726
Lys Ser Ile Gly Leu Glu Leu Gln Phe Ser Ile Pro Arg Leu Arg Gln			
135	140	145	
atc ggt ccg ggt gag agc tgc cca gac gga gtc act cac tcc atc agc			774
Ile Gly Pro Gly Glu Ser Cys Pro Asp Gly Val Thr His Ser Ile Ser			
150	155	160	
ggc cga atc gat gcc acc gtg gtc agg atc gga acc ttc tgc agc aat			822
Gly Arg Ile Asp Ala Thr Val Val Arg Ile Gly Thr Phe Cys Ser Asn			
165	170	175	180
ggc act gtg tcc cgg atc aag atg caa gaa gga gtg aaa atg gcc tta			870
Gly Thr Val Ser Arg Ile Lys Met Gln Glu Gly Val Lys Met Ala Leu			
185	190	195	
cac ctc cca tgg ttc cac ccc aga aat gtc tcc ggc ttc agc att gca			918
His Leu Pro Trp Phe His Pro Arg Asn Val Ser Gly Phe Ser Ile Ala			
200	205	210	
aac cgc tca tct ata aaa cgt ctg tgc atc atc gag tct gtg ttt gag			966
Asn Arg Ser Ser Ile Lys Arg Leu Cys Ile Ile Glu Ser Val Phe Glu			
215	220	225	
ggt gaa ggc tca gca acc ctg atg tct gcc aac tac cca gaa ggc ttc			1014
Gly Glu Gly Ser Ala Thr Leu Met Ser Ala Asn Tyr Pro Glu Gly Phe			
230	235	240	
cct gag gat gag ctc atg acg tgg cag ttt gtc gtt cct gca cac ctg			1062
Pro Glu Asp Glu Leu Met Thr Trp Gln Phe Val Val Pro Ala His Leu			
245	250	255	260
cgg gcc agc gtc tcc ttc ctc aac ttc aac ctc tcc aac tgt gag agg			1110
Arg Ala Ser Val Ser Phe Leu Asn Phe Asn Leu Ser Asn Cys Glu Arg			
265	270	275	
aag gag gag cgg gtt gaa tac tac atc ccg ggc tcc acc acc aac ccc			1158
Lys Glu Glu Arg Val Glu Tyr Tyr Ile Pro Gly Ser Thr Thr Asn Pro			
280	285	290	
gag gtg ttc aag ctg gag gac aag cag cct ggg aac atg gcg ggg aac			1206
Glu Val Phe Lys Leu Glu Asp Lys Gln Pro Gly Asn Met Ala Gly Asn			
295	300	305	
ttc aac ctc tct ctg caa ggc tgt gac caa gat gcc caa agt cca ggg			1254
Phe Asn Leu Ser Leu Gln Gly Cys Asp Gln Asp Ala Gln Ser Pro Gly			
310	315	320	
atc ctc cgg ctg cag ttc caa gtt ttg gtc caa cat cca caa aat gaa			1302
Ile Leu Arg Leu Gln Phe Gln Val Leu Val Gln His Pro Gln Asn Glu			
325	330	335	340

agc aat aaa atc tac gtg gtt gac ttg agt aat gag cga gcc atg tca	1350
Ser Asn Lys Ile Tyr Val Val Asp Leu Ser Asn Glu Arg Ala Met Ser	
345 350 355	
ctc acc atc gag cca cgg ccc gtc aaa cag agc cgc aag ttt gtc cct	1398
Leu Thr Ile Glu Pro Arg Pro Val Lys Gln Ser Arg Lys Phe Val Pro	
360 365 370	
ggc tgt ttc gtg tgt cta gaa tct cgg acc tgc agt agc aac ctc acc	1446
Gly Cys Phe Val Cys Leu Glu Ser Arg Thr Cys Ser Ser Asn Leu Thr	
375 380 385	
ctg aca tct ggc tcc aaa cac aaa atc tcc ttc ctt tgt gat gat ctg	1494
Leu Thr Ser Gly Ser Lys His Lys Ile Ser Phe Leu Cys Asp Asp Leu	
390 395 400	
aca cgt ctg tgg atg aat gtg gaa aaa acc ata agc tgc aca gac cac	1542
Thr Arg Leu Trp Met Asn Val Glu Lys Thr Ile Ser Cys Thr Asp His	
405 410 415 420	
cgg tac tgc caa agg aaa tcc tac tca ctc cag gtg ccc agt gac atc	1590
Arg Tyr Cys Gln Arg Lys Ser Tyr Ser Leu Gln Val Pro Ser Asp Ile	
425 430 435	
ctc cac ctg cct gtg gag ctg cat gac ttc tcc tgg aag ctg ctg gtg	1638
Leu His Leu Pro Val Glu Leu His Asp Phe Ser Trp Lys Leu Leu Val	
440 445 450	
ccc aag gac agg ctc agc ctg gtg ctg gtg cca gcc cag aag ctg cag	1686
Pro Lys Asp Arg Leu Ser Leu Val Leu Val Pro Ala Gln Lys Leu Gln	
455 460 465	
cag cat aca cac gag aag ccc tgc aac acc agc ttc agc tac ctc gtg	1734
Gln His Thr His Glu Lys Pro Cys Asn Thr Ser Phe Ser Tyr Leu Val	
470 475 480	
gcc agt gcc ata ccc agc cag gac ctg tac ttc ggc tcc ttc tgc ccg	1782
Ala Ser Ala Ile Pro Ser Gln Asp Leu Tyr Phe Gly Ser Phe Cys Pro	
485 490 495 500	
gga ggc tct atc aag cag atc cag gtg aag cag aac atc tcg gtg acc	1830
Gly Gly Ser Ile Lys Gln Ile Gln Val Lys Gln Asn Ile Ser Val Thr	
505 510 515	
ctt cgc acc ttt gcc ccc agc ttc caa caa gag gcc tcc agg cag ggt	1878
Leu Arg Thr Phe Ala Pro Ser Phe Gln Gln Glu Ala Ser Arg Gln Gly	
520 525 530	
ctg acg gtg tcc ttt ata cct tat ttc aaa gag gaa ggc gtt ttc acg	1926
Leu Thr Val Ser Phe Ile Pro Tyr Phe Lys Glu Glu Gly Val Phe Thr	
535 540 545	

gtg acc cct gac aca aaa agc aag gtc tac ctg agg acc ccc aac tgg	1974
Val Thr Pro Asp Thr Lys Ser Lys Val Tyr Leu Arg Thr Pro Asn Trp	
550 555 560	
gac cgg ggc ctg cca tcc ctc acc tct gtg tcc tgg aac atc agc gtg	2022
Asp Arg Gly Leu Pro Ser Leu Thr Ser Val Ser Trp Asn Ile Ser Val	
565 570 575 580	
ccc aga gac cag gtg gcc tgc ctg act ttc ttt aag gag cgg agc ggc	2070
Pro Arg Asp Gln Val Ala Cys Leu Thr Phe Phe Lys Glu Arg Ser Gly	
585 590 595	
gtg gtc tgc cag aca ggg cgc gca ttc atg atc atc cag gag cag cgg	2118
Val Val Cys Gln Thr Gly Arg Ala Phe Met Ile Ile Gln Glu Gln Arg	
600 605 610	
acc cgg gct gag gag atc ttc agc ctg gac gag gat gtg ctc ccc aag	2166
Thr Arg Ala Glu Glu Ile Phe Ser Leu Asp Glu Asp Val Leu Pro Lys	
615 620 625	
cca agc ttc cac cat cac agc ttc tgg gtc aac atc tct aac tgc agc	2214
Pro Ser Phe His His His Ser Phe Trp Val Asn Ile Ser Asn Cys Ser	
630 635 640	
ccc acg agc ggc aag cag cta gac ctg ctc ttc tcg gtg aca ctt acc	2262
Pro Thr Ser Gly Lys Gln Leu Asp Leu Leu Phe Ser Val Thr Leu Thr	
645 650 655 660	
cca agg act gtg gac ttg act gtc atc ctc atc gca gcg gtg gga ggt	2310
Pro Arg Thr Val Asp Leu Thr Val Ile Leu Ile Ala Ala Val Gly Gly	
665 670 675	
gga gtc tta ctg ctg tct gcc ctc ggg ctc atc att tgc tgt gtg aaa	2358
Gly Val Leu Leu Leu Ser Ala Leu Gly Leu Ile Ile Cys Cys Val Lys	
680 685 690	
aag aag aaa aag aag aca aac aag ggc ccc gct gtg ggt atc tac aat	2406
Lys Lys Lys Lys Lys Thr Asn Lys Gly Pro Ala Val Gly Ile Tyr Asn	
695 700 705	
ggc aac atc aat act gag atg ccg agg cag cca aaa aag ttt cag aaa	2454
Gly Asn Ile Asn Thr Glu Met Pro Arg Gln Pro Lys Lys Phe Gln Lys	
710 715 720	
ggg cga aag gac aat gac tcc cat gtg tat gca gtc atc gag gac acc	2502
Gly Arg Lys Asp Asn Asp Ser His Val Tyr Ala Val Ile Glu Asp Thr	
725 730 735 740	
atg gta tat ggg cat ctg cta cag gat tcc agc ggc tcc ttc ctg cag	2550
Met Val Tyr Gly His Leu Leu Gln Asp Ser Ser Gly Ser Phe Leu Gln	
745 750 755	

cca gag gtg gac acc tac cgg ccg ttc cag ggc acc atg ggg gtc tgt 2598
Pro Glu Val Asp Thr Tyr Arg Pro Phe Gln Gly Thr Met Gly Val Cys
760 765 770

cct ccc tcc cca ccc acc ata tgc tcc agg gcc cca act gca aag ttg 2646
Pro Pro Ser Pro Pro Thr Ile Cys Ser Arg Ala Pro Thr Ala Lys Leu
775 780 785

gcc act gag gag cca cct cct cgc tcc cct cct gag tct gag agt gaa 2694
Ala Thr Glu Glu Pro Pro Pro Arg Ser Pro Pro Glu Ser Glu Ser Glu
790 795 800

ccg tac acc ttc tcc cat ccc aac aat ggg gat gta agc agc aag gac 2742
Pro Tyr Thr Phe Ser His Pro Asn Asn Gly Asp Val Ser Ser Lys Asp
805 810 815 820

aca gac att ccc tta ctg aac act cag gag ccc atg gag cca gca gaa 2790
Thr Asp Ile Pro Leu Leu Asn Thr Gln Glu Pro Met Glu Pro Ala Glu
825 830 835

taa cttgatccat tccagacgct ttgctgagtt tcataaagca gggcactgag 2843

acacccggtcc gtgttcctaa ccagaaatcc taaagaagag gaattataca gaaggaacag 2903

caggagggttt tccctggacac cgccaacttc acattgctca gtggactcat tctaagggca 2963

agacattgaa aatgatgaat tccaatctgg atacagtcac gacagctcat gtgctcctca 3023

acttaggctg tgcgggttagc cagcctgtaa tgagaggaga gaggcctgag tcacctagca 3083

taggggttgca gcaagccctg gattcagagt gttaaacaga ggcttgccct cttcaggaca 3143

acagttccaa ttccaaggag cctacctgag gtcctactc tcactgggggt ccccaggatg 3203

aaaacgacaa tgtgcctttt tattattatt tatttggtgg tcctgtgtta tttaagagat 3263

caaagtata accacctagc tcttttcacc tgacttagta ataactcata ctaactgggt 3323

tggatgcctg gggtgtgact tctactgacc gctagataaa cgtgtgcctg tccccagggt 3383

gggtgggaata atttacaatc tgtccaacca gaaaagaatg tgtgtgtttg agcagcattg 3443

acacatatct gctttgataa gagacttcct gattctctag gtcgggtcgt gggtatocca 3503

ttgtggaaat tcactctgaa tccattgtc ctatagtcct agcaataaga gaaatttcct 3563

caagtttcca tgtgcggttc tcttagctgc agcaataactt tgacatttaa agagaaattt 3623

agagaatatt ctcactcctc aaaaatgttt aatatatac caaacagtgg cccctgcat 3683

tagttttctg ttgccactgc aaccattac ttggtagctt aaaaacaaca cattagctta 3743

tagtcctggg gatcagaatt ccaaaatgga tgtccctgaa tgaaaatcaa ggtgtcagca 3803
gagctgtgct ccttctgaag gctctagga gaagccggtt ccttgccatt tcaagcttct 3863
agaggctggc tgcattccca ggctccagt gctgggtcaag cttttctcac atggcatcac 3923
tgtgacactg gccctccac ttccctcttt gacttacaaa gccaccagg aagatccagg 3983
ataatctctc catctaaaga tccttcatca tcctggaaga gccttttgcc atgcaagaca 4043
acatagccac aggtggggat taggaccagg acatctttgg ggtgctgtta ttctgcctac 4103
cacaccttcc tgccactgac tcccacagga gaggctacaa aatgatctgg cgcacaggga 4163
tgttttgttt agcttgcgga ctctaact taaaaaacc ccagatcaga agatctggcc 4223
atgctggggc tcacattctc acctagcaac aactggctgg agctgggcac cagctctgcc 4283
tttagaagg ggtgccactt caccaggta ccacagcca cactacgcc tatcacttcc 4343
cacaatgagg ctaagtgttt gtttctactg atcaatgcc ctgcagggtg catttattgt 4403
aatgaaaaag aaagactggg attaatctct aatcagggtg gtagaccatg agaccaatgt 4463
gtgctcacat tacccttttt cttttttttt tttttctttt tctttttttt tttaatgtga 4523
gacaggatct cattctgttg cctaggctgg agtgcagtgg cgcaatctcg gctcactgca 4583
acctctgcct cctgggctca agcaattctc ccacctcagc ctcccaaata gctgggatca 4643
ctggcacaaa ccaccatgcc cagctaattt tgtatttttt gtagagacag ggtttcacca 4703
tgttgcccag gctgggtctca acctcctggg ctcaagcaat cctcctgcct cggcctccca 4763
aagtgtggg attacagatg tgagccaccg catccagccc cacaccctca ttataccaa 4823
ttacctgcc agtaactgtg gacttttgct tcctcaccct tgctctgatc tggaaggaga 4883
gggattatgt tatagcttgt cagcacagtc ccaagttcaa tatttctgcg gcaaaaactt 4943
ccttcaaaaa ataaatgtac ttcatgtat tcaatgaatt caccttgga atgcaccgcc 5003
tcaacttggt cacatggcat aaatgaaagg aattttatag tctcctaaat ggcgtgtact 5063
gcaagacctc ttgaacactt tccagaggat aggatattta agtcatgccc ttggcgttgc 5123
ctatggcacc ttcccttct gaaagtctgg ttctgccc gtgacccttg gccttgtag 5183
ccgagatgct gaccctgcat aaagggccaa aggagggtg cggcttcctt cctcactga 5243
agagccctta ttgaattca ctgtgtggag ccctagccct ccattctcga cattcccaa 5303
cctcccagcc cttccaagc aggactaggt gccctgcatt ccaccaagg tgggattggc 5363


```

cttccttagg ctggctactt gtcaccatca ccgacatcac tgttgctgc aaggacacca 5423
cgtggccatt ttccttcaac tgagggctca aaactcctgg acaagttgct ggctcctgag 5483
accagtattt cctggagctg tgcctcagtg aaggggcca gcctgaggaa ccctggctct 5543
tttctttaaa gcccaggccc cacttacata aaacatttca gggtcactgg aaacagtgaa 5603
gtgccatttg ttgaagccta ctgcatgcca gccactgct catccacgtg gtctgccatg 5663
cctacgagga aggccagcgc atgcaggact ggtctcta at gctgtggtca ttgcacagaa 5723
gggaaagggtc tcaaggaaga gtcaactggg acaagcaciaa gccaccgga catggccttg 5783
gtaaagggtta gcagactggt gtgtgtggat ctgcagtgtc tcaactggaaa taatttattc 5843
attgcagata ctttttaggt ggcattttat tcatttcttg tgctttaaat aaacaaatgt 5903
acaaaaaac aagtatcaag ctgtttaagt gcttcggcta cttgtccctt gggtcagtag 5963
aggccccggt ttcccagttg ttgactgtga caggctcagc atgggctcag cagatgctgt 6023
cttaatttgt ggatgatata gaaagccagg ctttgggata caagttcttt cctcttcatt 6083
tgatgccgtg cactgtgtga agcagatggt tttgtccgga aataaaaata atagtcttgg 6143
agtctcgcca aaaaaaaaaa 6163

```

<210> 4
 <211> 836
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Ala Gly Leu Asn Cys Gly Val Ser Ile Ala Leu Leu Gly Val Leu
 1 5 10 15
 Leu Leu Gly Ala Ala Arg Leu Pro Arg Gly Ala Glu Ala Phe Glu Ile
 20 25 30
 Ala Leu Pro Arg Glu Ser Asn Ile Thr Val Leu Ile Lys Leu Gly Thr
 35 40 45
 Pro Thr Leu Leu Ala Lys Pro Cys Tyr Ile Val Ile Ser Lys Arg His
 50 55 60
 Ile Thr Met Leu Ser Ile Lys Ser Gly Glu Arg Ile Val Phe Thr Phe
 65 70 75 80
 Ser Cys Gln Ser Pro Glu Asn His Phe Val Ile Glu Ile Gln Lys Asn
 85 90 95

Ile	Asp	Cys	Met	Ser	Gly	Pro	Cys	Pro	Phe	Gly	Glu	Val	Gln	Leu	Gln		
			100					105					110				
Pro	Ser	Thr	Ser	Leu	Leu	Pro	Thr	Leu	Asn	Arg	Thr	Phe	Ile	Trp	Asp		
			115				120					125					
Val	Lys	Ala	His	Lys	Ser	Ile	Gly	Leu	Glu	Leu	Gln	Phe	Ser	Ile	Pro		
	130					135					140						
Arg	Leu	Arg	Gln	Ile	Gly	Pro	Gly	Glu	Ser	Cys	Pro	Asp	Gly	Val	Thr		
145					150					155					160		
His	Ser	Ile	Ser	Gly	Arg	Ile	Asp	Ala	Thr	Val	Val	Arg	Ile	Gly	Thr		
				165					170					175			
Phe	Cys	Ser	Asn	Gly	Thr	Val	Ser	Arg	Ile	Lys	Met	Gln	Glu	Gly	Val		
			180					185					190				
Lys	Met	Ala	Leu	His	Leu	Pro	Trp	Phe	His	Pro	Arg	Asn	Val	Ser	Gly		
	195					200						205					
Phe	Ser	Ile	Ala	Asn	Arg	Ser	Ser	Ile	Lys	Arg	Leu	Cys	Ile	Ile	Glu		
	210					215					220						
Ser	Val	Phe	Glu	Gly	Glu	Gly	Ser	Ala	Thr	Leu	Met	Ser	Ala	Asn	Tyr		
225					230					235					240		
Pro	Glu	Gly	Phe	Pro	Glu	Asp	Glu	Leu	Met	Thr	Trp	Gln	Phe	Val	Val		
				245				250						255			
Pro	Ala	His	Leu	Arg	Ala	Ser	Val	Ser	Phe	Leu	Asn	Phe	Asn	Leu	Ser		
			260					265					270				
Asn	Cys	Glu	Arg	Lys	Glu	Glu	Arg	Val	Glu	Tyr	Tyr	Ile	Pro	Gly	Ser		
	275						280					285					
Thr	Thr	Asn	Pro	Glu	Val	Phe	Lys	Leu	Glu	Asp	Lys	Gln	Pro	Gly	Asn		
	290					295					300						
Met	Ala	Gly	Asn	Phe	Asn	Leu	Ser	Leu	Gln	Gly	Cys	Asp	Gln	Asp	Ala		
305					310					315					320		
Gln	Ser	Pro	Gly	Ile	Leu	Arg	Leu	Gln	Phe	Gln	Val	Leu	Val	Gln	His		
				325					330					335			
Pro	Gln	Asn	Glu	Ser	Asn	Lys	Ile	Tyr	Val	Val	Asp	Leu	Ser	Asn	Glu		
			340					345					350				
Arg	Ala	Met	Ser	Leu	Thr	Ile	Glu	Pro	Arg	Pro	Val	Lys	Gln	Ser	Arg		
	355						360					365					
Lys	Phe	Val	Pro	Gly	Cys	Phe	Val	Cys	Leu	Glu	Ser	Arg	Thr	Cys	Ser		

370		375		380
Ser Asn Leu Thr Leu Thr Ser Gly Ser Lys His Lys Ile Ser Phe Leu				
385		390		395
				400
Cys Asp Asp Leu Thr Arg Leu Trp Met Asn Val Glu Lys Thr Ile Ser				
		405		410
				415
Cys Thr Asp His Arg Tyr Cys Gln Arg Lys Ser Tyr Ser Leu Gln Val				
		420		425
				430
Pro Ser Asp Ile Leu His Leu Pro Val Glu Leu His Asp Phe Ser Trp				
		435		440
				445
Lys Leu Leu Val Pro Lys Asp Arg Leu Ser Leu Val Leu Val Pro Ala				
		450		455
				460
Gln Lys Leu Gln Gln His Thr His Glu Lys Pro Cys Asn Thr Ser Phe				
		465		470
				475
Ser Tyr Leu Val Ala Ser Ala Ile Pro Ser Gln Asp Leu Tyr Phe Gly				
		485		490
				495
Ser Phe Cys Pro Gly Gly Ser Ile Lys Gln Ile Gln Val Lys Gln Asn				
		500		505
				510
Ile Ser Val Thr Leu Arg Thr Phe Ala Pro Ser Phe Gln Gln Glu Ala				
		515		520
				525
Ser Arg Gln Gly Leu Thr Val Ser Phe Ile Pro Tyr Phe Lys Glu Glu				
		530		535
				540
Gly Val Phe Thr Val Thr Pro Asp Thr Lys Ser Lys Val Tyr Leu Arg				
		545		550
				555
Thr Pro Asn Trp Asp Arg Gly Leu Pro Ser Leu Thr Ser Val Ser Trp				
		565		570
				575
Asn Ile Ser Val Pro Arg Asp Gln Val Ala Cys Leu Thr Phe Phe Lys				
		580		585
				590
Glu Arg Ser Gly Val Val Cys Gln Thr Gly Arg Ala Phe Met Ile Ile				
		595		600
				605
Gln Glu Gln Arg Thr Arg Ala Glu Glu Ile Phe Ser Leu Asp Glu Asp				
		610		615
				620
Val Leu Pro Lys Pro Ser Phe His His His Ser Phe Trp Val Asn Ile				
		625		630
				635
Ser Asn Cys Ser Pro Thr Ser Gly Lys Gln Leu Asp Leu Leu Phe Ser				
		645		650
				655

Val Thr Leu Thr Pro Arg Thr Val Asp Leu Thr Val Ile Leu Ile Ala
660 665 670

Ala Val Gly Gly Gly Val Leu Leu Leu Ser Ala Leu Gly Leu Ile Ile
675 680 685

Cys Cys Val Lys Lys Lys Lys Lys Lys Thr Asn Lys Gly Pro Ala Val
690 695 700

Gly Ile Tyr Asn Gly Asn Ile Asn Thr Glu Met Pro Arg Gln Pro Lys
705 710 715 720

Lys Phe Gln Lys Gly Arg Lys Asp Asn Asp Ser His Val Tyr Ala Val
725 730 735

Ile Glu Asp Thr Met Val Tyr Gly His Leu Leu Gln Asp Ser Ser Gly
740 745 750

Ser Phe Leu Gln Pro Glu Val Asp Thr Tyr Arg Pro Phe Gln Gly Thr
755 760 765

Met Gly Val Cys Pro Pro Ser Pro Pro Thr Ile Cys Ser Arg Ala Pro
770 775 780

Thr Ala Lys Leu Ala Thr Glu Glu Pro Pro Pro Arg Ser Pro Pro Glu
785 790 795 800

Ser Glu Ser Glu Pro Tyr Thr Phe Ser His Pro Asn Asn Gly Asp Val
805 810 815

Ser Ser Lys Asp Thr Asp Ile Pro Leu Leu Asn Thr Gln Glu Pro Met
820 825 830

Glu Pro Ala Glu
835

<210> 5

<211> 23

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 5

accgcctcaa cttgttcaca tgg

<210> 6

<211> 26

<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 6
ctgggtctcag gagccagcaa cttgtc 26

<210> 7
<211> 25
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 7
ctcatgacgt ggcagtttgt cgttc 25

<210> 8
<211> 26
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 8
ggctcgctca ttactcaagt caacca 26

<210> 9
<211> 36
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 9
attcgcgact gatgatcgat tttttttttt tttttt 36

<210> 10
<211> 20
<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 10

attcgcgact gatgatcgat

20

<210> 11

<211> 20

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 11

gagatattag aattctactc

20

<210> 12

<211> 17

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 12

gagtagaatt ctaatat

17

<210> 13

<211> 22

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 13

agtccatgtg aacaagttga gg

22

<210> 14

<211> 20

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 14

aattctccca cctcagcctc

20

<210> 15

<211> 22

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 15

aggatgaaaa cgacaatgtg cc

22

<210> 16

<211> 21

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 16

agaattgctt gagcccagga g

21

<210> 17

<211> 21

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 17

caacttcaca ttgctcagtg g

21

<210> 18

<211> 25

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 18
tgagcaagtt cagcctgggtt aagtc 25

<210> 19
<211> 26
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 19
caccgaatac tcataaagaa ggtccc 26

<210> 20
<211> 26
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 20
tagacttcga gcaggagatg gccact 26

<210> 21
<211> 20
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 21
ccagccatgt acgtagccat 20

<210> 22
<211> 19
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 22
ccaagaagga aggctggaa 19

<210> 23
<211> 25
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 23
ccatcaccat cttccaggag cgaga 25

<210> 24
<211> 19
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 24
ccaagaagga aggctggaa 19

<210> 25
<211> 20
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 25
tgcaggaggc attgctgatg 20

<210> 26
<211> 19
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 26
aaatcgtgca cttgcaggc 19

<210> 27
<211> 18
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 27
ttgatgcgtt ccagctga 18

<210> 28
<211> 21
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 28
ttgaattcac tgtgtggagc c 21

<210> 29
<211> 19
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 29
tgcaggcaac agtgatgtc 19

<210> 30
<211> 24
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 30
attggccttc cttaggctgg ctac 24

<210> 31
<211> 43

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 31

tgtagcgtga agacgacaga aagggcgtgg taccgagctc gag

43

<210> 32

<211> 22

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 32

agggcgtggt accgagctcg ag

22

<210> 33

<211> 11

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 33

ggctcgagct c

11

<210> 34

<211> 22

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 34

ggccatgtcc ggtgggcttg tg

22

<210> 35

<211> 26

<212> DNA

<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 35
ctcaaaactc ctggacaagt tgctgg 26

<210> 36
<211> 22
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 36
aaggtgaagg tcggagtcaa cg 22

<210> 37
<211> 24
<212> DNA
<213> Kunstliche Sequenz

<220>
<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 37
ggcagagatg atgacccttt tggc 24

<210> 38
<211> 23
<212> DNA
<213> Kunstliche Sequenz

<220>
<221> 5'UTR
<222> (1)..(282)

<220>
<221> GC_signal
<222> (147)..(157)

<220>
<221> misc_feature
<222> (201)..(209)
<223> cap signal; Transkriptionsstart

<220>
<221> 3'UTR

$\langle 220 \rangle$ -

<222> (2794) .. (6163)

<220>

<221> CDS

<222> (283) .. (2793)

<400> 38

agcagcagaa cccctagcag tgc

23

<210> 39

<211> 26

<212> DNA

<213> Kunstliche Sequenz

<220>

<223> Beschreibung der kunstlichen Sequenz: Primer

<400> 39

agaaccccta gcagtgcgat agagac

26

<210> 40

<211> 27

<212> DNA

<213> Kunstliche Sequenz

 $\langle 220 \rangle$

<223> Beschreibung der künstlichen Sequenz: Primer

<400> 40

gaactgtaat gttgctttct cgtggca

27